DOCUMENT RESUME

ED 216 030 TM 820 184

AUTHOR Stufflebeam, Daniel L.
TITLE Reflections on the Movement to Promote Effective Educational Evaluations Through the Use of Professional Standards. Discussion Draft.
PUB DATE Aug 81
NOTE 33p.; Paper presented at the Annual Meeting of the American Psychological Association (Los Angeles, CA, August, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Curriculum Evaluation; *Evaluation Methods; Program Evaluation; *Standards
IDENTIFIERS *Standards for Evaluation Educ Prog Proj Materials

ABSTRACT

On behalf of the Joint Committee on Standards for Educational Evaluation, the author reviews the Committee's recent efforts and discusses the needs for further work related to the "Standards for Evaluations of Educational Programs, Projects, and Materials." The Committee has developed a program to guide its work. The program addresses four areas of activity: provision of training and technical assistance; research, development, and training for interpretation and use of the Standards; dissemination of information related to the standards; and revision of the Standards as needed. Critiques of the Standards appeared in the May 1981 issue of "Evaluation News." A content analysis of these critiques yielded both positive and negative comments in four areas: content, rationale, structure, and use. Further work on the content of the Standards should address the trade-offs among them, the use of needs assessment in evaluation, and audiovisual training materials. Dialogue among various standard-setting groups would further develop the Standards' rationale. The structure of the standard-setting process could be improved by expanding the Committee membership. The use of the Standards in field tests should be encouraged. (BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED216030

Discussion Draft

Reflections on the Movement to Promote Effective Educational Evaluations Through the Use of Professional Standards

Presented by:

Daniel L. Stufflebeam
Chairman of the Joint Committee on Standards
for Educational Evaluation

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Stufflebeam, D.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

at

the Annual Meeting of the American Psychological Association

Los Angeles, California

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

August, 1981

TM 820 181

I am speaking to you today on behalf of the Joint Committee on Standards for Educational Evaluation. As you are aware, the Committee, last year, after nearly 5 years of work, issued Standards for Evaluations of Educational Programs, Projects, and Materials. The Standards, as I will refer to this publication hereafter in this report, is intended to provide evaluators and their clients with an authoritative guide for assessing and upgrading the quality of program evaluations in education.

At last year's APA meeting in Montreal, I reviewed the work involved in developing the Standards, and I described the Committee's plans for an ongoing program to promote the effective use of the Standards and periodically to accomplish needed revisions. This year I have been invited to bring you up-to-date on the Committee's recent efforts and specifically to discuss the needs for further work related to the Standards. I am pleased to respond to Dr. Diamond's invitation, because the Committee wants to continue to be accountable to its Sponsoring Organizations, and because, as a volunteer group we need to involve as many interested persons as possible in our work.

I have organized the body of my report into three parts. In the first part, I will review pertinent developments during this past year. Then I will identify and analyze issues concerning the Standards that I have drawn from a series of critiques of the Standards which appeared recently in Evaluation News. Finally, I will attempt to address the identified issues by proposing a partial agenda for conducting projects designed to improve both the contents and uses of the Standards. Without further introduction, I turn to my brief progress report.

1. Progress This Past Year

The Standards has been quite well received. At last report about 4,000 copies had been sold and McGraw-Hill was in the process of ordering a second printing.

The Committee and its Sponsoring Organizations have taken several steps to meet the needs for an ongoing standard-setting effort. The Committee developed a plan for a standing committee, developed a set of Principles and By-laws to guide and govern the work of the new Committee, and became a legally constituted body.

Eleven organizations decided to sponsor the continuing effort; and each one appointed a representative (See Exhibit 1) to serve on the Committee and agreed to cover the member's expenses associated with the Committee's annual meeting. The new representative from APA is Dr. Carol Kehr Tittle. The Committee arranged to use the royalties from the sales of its publications to help support its work; while minimal, these funds provide a measure of ongoing support. The Committee also obtained grants totalling about $6,000 from the EXXON Foundation and the International Paper Company Foundation to support basic planning and organizing activities. Currently, the Committee is developing funding requests for a number of specific projects.

The Committee has also begun to respond to a number of inquiries and requests for assistance. For example, individual members have reacted to the efforts of personnel of the Louisiana Department of Education to apply the Standards on a statewide basis; members have conducted workshops on the Standards in Malaysia, Australia, and New Zealand, as well as the U.S., and personnel of the Evaluation Center have responded to numerous requests for information. Also The Committee has collaborated with the ERIC Center at ETS toward the development and publication of an annotated bibliography to accompany the Standards.

In addition to the preceding specific steps, the Committee has defined a general program to guide its work. This program is divided into the following four areas:

1. Provision of training and technical assistance to evaluators and users of evaluations.
2. Research, development, and training associated with interpreting and using the Standards.
3. Dissemination of and clearinghouse for information related to the Standards and their use; and
4. Revision of the Standards as needed, and expansion of their use in new directions (e.g., evaluations of personnel and facilities).

The preceding report of activity demonstrates that the Committee has laid a substantial foundation for an ongoing effort to develop and promote sound use of evaluation standards. Next, the Committee must begin implementing their program by choosing and carrying out specific projects that respond to high priority needs.

2. Analysis of Issues From the Literature

One basis for identifying these needs is to be found in the May 1981 issue of Evaluation News. That issue contains a critical appraisal of the Standards. The 10 articles that convey this appraisal include the following: general reviews by Sechrest and Stake; implications of the Standards for licensure and accreditation by Becker and Kirkhart; Federal and State perspectives, respectively, by Wargo and Baron; an urban school district perspective by Osterlind; a rural school perspective by Hecht; an out-of-school learning perspective by Marcia Linn; an international perspective by Searle; and a mental health programs perspective by Lyons and Rubin. Collectively, these articles provide a rich and varied view of the Standards.

The Joint Committee has always believed that one of its main responsibilities is to seek out, study, and react to critical appraisals and field tests of its products. Such evidence provides a vital basis for ensuring that the Committee's work is sensitive and responsive to what persons in the field see as real and important evaluation problems. It is fitting, then, that the Joint Committee seriously study the reactions to their work that appeared in the May, 1981 issue of Evaluation News.

In order to mine the articles for potentially useful leads, I performed a content analysis across all eleven of them. First, I made a list of allegations about the Standards from each article (See Exhibit 2). In accordance with the views presented by the authors, I divided the allegations into lists of strengths, weaknesses, and suggestions. Then I sorted the items in each list into four categories that were previously found to be useful for organizing research and development activities in relation to the standards for accounting and auditing.

As reported by Ridings[1], these categories are: 1) the rationale for stating and using standards in a field; 2) the content of the standards; 3) the structure by which the standards are set, reviewed, and modified; and 4) uses of the standards. The result of these sorts was a matrix of strengths, weaknesses, and recommendations in each of the four categories (See Exhibit 3).

This matrix revealed a number of general characteristics of the eleven reviews. On the whole, the critics found the Standards to be responsive to a real need and of high quality. In a number of respects, however, they pointed to needs for further testing and revision. Occasionally, what one critic saw as a strength was viewed by another as a weakness; for example, Wargo applauded the specific advice in Standard A1 regarding rank ordering of audiences, while Stake cited this an an example of overspecification. Generally, it seemed clear that all of the critiques were armchair evaluations, since none of them identified any strengths and weaknesses as regards actual uses of the Standards. But a number of the authors did recommend that the Standards be field tested in a wide range of settings. Clearly, the observations made by these authors should be viewed as tentative and subject to verification, and not a sufficient guide for revising the Standards and the standard setting process.

Nevertheless, the matrix of information does point up a number of problems that might be addressed by the Committee or other groups. I shall mention a few of these to give you a general idea of the needs identified by the critics. I'll discuss each of the four categories in turn.

The most comments were contained in the category which concerns the content of the Standards. These comments reflected extensive praise of the

---

[1] Ridings, Jeri M. Standard Setting in Accounting and Auditing: Considerations for Educational Evaluation. Kalamazoo, Michigan: Western Michigan University, 1980 (unpublished dissertation).

document. Among the positive statements were the following:

--the advice in the Standards is realistic, practical, usually of admirable temperament, and commending of adherence.

--the Standards is aimed at a varied audience and provides a comprehensive, useable, and "concise" statement of good practice.

--the scope of the Standards is thorough, and the document

- raises a large number of key issues,
- offers many good suggestions,
- endorces methodological variety,
- focuses on a variety of evaluation situations,
- provides a range of interesting and realistic illustrations; and
- confronts a range of political realities.

Despite the overall positive assessment of the content of the Standards, the authors also offered a number of criticisms. Several of the writers pointed to the problem of conflicts between different standards and observed that the Committee hadn't offered enough useful advice about how to identify and address problems involving trade-offs among the standards. Whereas there were many compliments about the scope of the Standards, the document was also criticized for being less than comprehensive. Stake, for example, noted that the Standards is incomplete in its identification of instances of malpractice that are common in the educational evaluation establishment, such as promising what can't be done. Several writers also observed or implied that the Standards should explicitly require that needs assessment be a part of every evaluation. Sechrist criticized the Committee for not including a standard requiring the best applicable research design. The essence of these criticisms is that the scope of topics treated and practices required should be expanded.

Other criticisms concerned the useability of the Standards. They were seen not to reflect considerations found in rural settings, Federal evaluations, and evaluations in foreign countries.

In general, the criticisms of the contents of the Standards was constructive and pointed to possibilities for improvements. This is borne out by the many recommendations that the reviewers offered. They called for companion volumes to help various groups use the Standards. They suggested adding new standards and other illuminating material in the next edition of the Standards. And they suggested how several of the present 30 standards might be modified. I have made an inventory of these suggestions and will present them for consideration and further study at the next meeting of the Joint Committee in October.

Turning from the content of the Standards, the eleven articles I examined also provided insights concerning the rationale for an effort to set and use standards in educational evaluation. As Ridings pointed out, it is fundamentally important that standard setters periodically review the role of standards in their profession in order to ensure that positive ends are being served and that negative side effects are being minimized. Stake granted that the rationale for standards has not been put better than by the Joint Committee. Other writers observed that the Standards likely would have many positive benefits.

These include providing

--benchmarks for judging evaluations

--content for inservice and preservice training

--a guide for developing and licensing training programs

--a guide for developing standardized tests on evaluation

--a framework for a competency-based approach to accrediting and licensing evaluators

--a framework for developing evaluation contracts

--a stimulous for better conduct and use of evaluation; and

--encouragement for improving the integrity of evaluators

While some of these purposes might be controversial, on the whole they present a positive view of the potential effects of standards for evaluation.

On the negative side of the ledger, Stake saw some problems in how the Committee had addressed the issue of a rationale for standards in evaluation. He said that whereas the Committee had made a strong case for standards, they hadn't been sufficiently vigorous in examining the case against standards. He said that standards must certainly constrain creativity, and he speculated that they might unjustly deny to individuals the privilege to practice evaluation.

If I may be permitted to editorialize on this point, I think the Joint Committee would have me emphasize that they see the Standards serving more to promote and aid quality evaluation than to constrain or punish evaluators. Nevertheless, I must acknowledge that at least one department of education is attempting to use training in the Standards as one basis for granting the right to practice educational evaluation. This comes close to illustrating what Stake sees as a potential debilitating effect of the Standards. Since otherwise highly qualified evaluators might be prevented from practicing evaluation in the state simply because they hadn't attended the designated training session. Clearly the Joint Committee must be vigilant and proactive in this leadership towards helping the field to make sound use of the Standards and avoid some of the negative possibilities.

Another issue regarding the rationale for standards was pointed up in the article by Lyons and Rubin. They referenced about half a dozen standard setting efforts that have relevance to educational evaluation. Ridings and other wirters have pointed out that multiple standard setting programs in a given field create competing efforts that have the potential to confuse clients, waste resources, and fragment professional development. However, Lyons and

Rubin argued that evaluators' needs in different contexts are quite different, and they saw the creation of multiple sets of standards as potentially more responsive to idiosyncratic needs than would be the case with a single set of general standards. Stake agreed with this position when he called for different sets of standards that reflect local norms and conditions.

On this point, it seems clear that the Joint Committee and other standards setting groups must carefully consider how they can collectively and individually best serve the needs of education. At the very least the Joint Committee and the body responsible for the ERS standards should increase their dialogue about possibilities for collaboration.

This point is a good one for a transition to the matter of the adequacy of the Joint Committee's structure for setting standards. Given that standards for evaluations are needed, and given that their acceptability to the field depends on the credibility of the process by which they were developed, then it is important that the Joint Committee periodically review and, where necessary, revise the structure of their standard setting process.

In general, the authors of the articles on the Standards which appeared in Evaluation News observed that the developers of the Standards possessed high credibility. They also acknowledged that the Joint Committee made good use of a large and diverse number of advisors.

However, several of the writers pointed to deficiencies in this area. Especially, they said that the developers and support groups underrepresented certain groups. These included federal agency personnel, contract research organizations, rural school personnel, and foreign groups. The general point is well taken. I believe the Joint Committee continuously must attempt to ensure that all stakeholder groups are adequately represented in the standard setting effort.

How to do this leads us to our final area of concern regarding the Standards; this is the matter of Uses, and I ought to add, Misuses of the Standards. I've already noted that none of the eleven articles being discussed made any reference to actual uses of the Standards. However, they did offer recommendations in this realm. Wargo emphasized, for example, that the standards should be field tested at the Federal level, as a basis for revising them, so that they would be more useful to federal level evaluators--Marcia Linn suggested that the standards be field tested in out-of-school settings and in a variety of other "very different" settings such as other countries; she thought such tests would be very instructive as regards limitations and needs for revisions in the present Standards. Hecht added that the Committee should respond to the needs of rural school personnel by disseminating the Standards through non-print media. Finally, Stake implied that the Committee has a responsibility to help the field avoid uses of the Standards that constrain individual practice.

The preceding comments about all four aspects of the Standards being considered--i.e., content, rationale, structure, and use--have direct relevance to the future work of the Joint Committee. Overall, these comments indicate that the work of the Committee, to this point, is appreciated, respected, but not yet complete. These are recommendations for improvements and extensions, and they underscore the importance of the Committee's future role as regards improvement of the contents and uses of the Standards. Also there is clearly a need to provide strong leadership to ensure that the standard setting effort will continue to serve worthy purposes. There is a need to increase the effort to ensure that all stakeholder groups are properly involved and served. And there is definitely a need to obtain empirical evidence about uses of the standards.

3. A Proposed Agenda of Needed Work

How to address these four areas of need is the topic with which I will close this paper. Specifically, I will propose an agenda of activities. This agenda is intended partially for the consideration of the Joint Committee as they plan and seek funds to support their work. But it is also intended for researchers and developers who are interested in the area of evaluation.

The agenda that I would have us consider directly reflects the preceding analysis. I have identified what I consider to be one or more high priority projects in each of the four areas--content, rationale, structure, and use. In selecting these projects I have attempted to respond to the information contained in the eleven Evaluation News articles. The proposed projects mainly reflect an action research agenda.

In the area of content, I have three projects in mind. These are: 1) identifying and dealing with trade-offs among the standards; 2) developing and testing a standard that requires the use of needs assessment in evaluation; and 3) developing audio visual materials to assist trainers in training non-technically-oriented audiences to understand and use the Standards.

Regarding the rationale for Evaluation Standards, I believe the time has come for dialogue among the various groups which are involved in setting standards for evaluation. A well planned conference involving the representatives of these groups would clearly be in order. The conference could address questions concerning the relative merit of unified versus diversified standard-setting efforts. It could also search out and promote areas for future collaboration. I think the participants in such a conference should represent the ERS and Joint Committee standard setting groups, as well as other relevant groups, such as the Committee on Standards for Educational and Psychological tests. Moreover, the conference group definitely should cut across national boundaries.

In the third work area--that is concerned with the structure of the process--I believe the Joint Committee must take concrete steps to respond to the criticisms that were offered by Wargo, Searle, and Hecht. Consideration should be given to expanding the membership of the Joint Committee, so as better to reflect perspectives that are not represented now. Also the Committee should consider other means of increasing communication such as review panels, and a newsletter.

Finally, as regards use of the Standards I believe the Committee should encourage field tests of the standards in a variety of settings. Especially they should promote the conduct of case studies in other countries, at the Federal level in our country, in contract research corporations, in rural schools, and in out-of-school settings. The Committee should arrange to use the case reports in revising the Standards and should publish them in a form designed to help users of the Standards to see them in contexts that have particular meaning to them.

The projects I have just referenced, of course, are incompletely described and must be deliberated by the Joint Committee and its constituents. In presenting them I have tried to show the kinds of work that might be done to improve the content and use of the standards for evaluations. I have also tried to illustrate that feedback from the profession is a vital means of determining what needs to be done.

Exhibit 1

THE JOINT COMMITTEE

(As of July 1, 1981)

Esther Diamond (Science Research Associates), representing the Association for Measurement and Evaluation in Guidance

Roy Forbes (Education Commission of the States), representing the Education Commission of the States

Freda Holley (Austin, Texas Independent School District), representing the American Educational Research Association

Philip Hosford (New Mexico State University), representing the Association for Supervision and Curriculum Development

William Mays, Jr. (Michigan Elementary and Middle Schools Principals Association), representing the National Association for Elementary School Principals

Bernard McKenna (National Education Association), representing the National Education Association

Lloyd Nielsen (Roseville, Minnesota School District), representing the American Association of School Administrators

James Oglesby (Columbia, Missouri Public Schools), representing the National School Boards Association

James Sanders (Western Michigan University), representing the Evaluation Network

Daniel Stufflebeam (Western Michigan University), representing the National Council on Measurement in Education

Carol Kehr Tittle (University of North Carolina at Greensboro), representing the American Psychological Association

Exhibit 2

Critical Reactions to the Standards

	Pros	Cons	Suggestions
Lee Sechrest	--Benchmark for judging evaluations --Content for training --Useful inventory of issues --Credible developers --Consistent format which eases use and amendment --Realistic and commanding of adherence --Interesting and realistic illustrations --Potentially useful for contracting.	--Number (too many to live up to) --Incompatibilities (trade-offs) --Despite advice in Audience Identification. Sometimes it is legitimate to ignore an audience --Lack of specific reference to appropriate experimental designs.	--Add a standard calling for only valid and dependable designs. (best possible research design)
Robert Stake	--Case for standards has best been stated by the Committee (p. 149) --Standards that can't be measured are okay, because they arouse people to care about what someone admires --Standards raise a large number of important issues --Most of the statements are of admirable temperament.	--Committee didn't vigorously examine the case against the standards. --Standards bring injustice and constraints on creativity --Standards will aid unjust denials of the privilege to practice evaluation. [We already see this in Lousisana].	--Avoid constraining individual practice if done only to protect our professional way of life.

Exhibit 2 (continued)

	Pros	Cons	Suggestions
Robert Stake (continued)		--Some of the standards (e.g., audience identification are over-specified) --Standards haven't identified many bad practices (by "QUANGO") so tacitly legitimate them (p. 152)	--Implies a need for standards that prohibit unrealistic promise, being indifferent to concerns of evaluees, and promote efforts to understand education. --Provide standards by examples in our meta-evaluations as an alternative to formal standards.
Heather Becker and Karen Kirkhart	--The Standards provide a frame work for probing questions pertaining to developing a competency-based approach to accrediting and licensing evaluators. - Provide a possible guide for developing and accrediting training programs - Provide a possible basis for developing a standardized test for assessing one's knowledge of evaluation models, techniques, and principles - Guidelines could be used to develop behavioral criteria for judging an evaluator's performance.		

Exhibit 2 (continued)

	Pros	Cons	Suggestions
Heather Becker and Karen Kirkhart	--Scope of the Standards is most thorough (and are sufficient in breadth to serve as a basis for licence (accreditation).	--Exclusion of needs assessment --But they don't help with the choice of those that are most relevant for licence and accreditation	--Propriety and accuracy are viewed as most important in accreditation and licensing --Suggests possibility of a Consumer's Guide to Program Evaluation as a companion to the Standards. --The Standards require fine tuning, discussion, and negotiation before the Standards could be used for licensing based on criteria measures. (many issues must be resolved)
Michael J. Wargo	--Provide a comprehensive yet consise statement of "good practices" --Good organization and format --Functional reorganization useful --Endorcement of methodological variety --Emphasis on evaluation quality --Rationale for standards --Focus (small-large, etc.) --Confronts political realities --Targeting at varied audience --Agrees with need to rank order audiences	--Underrepresents Federal level groups on Jt. Comm. and support groups --Underrepresents contract research organizations --Unrealistic to assume evaluators are sufficiently independent to be able to quit --Standards fail to recognize and deal with the complexities of team work involved in federal evaluations --Most case studies are at state and local level --But says information needs within audiences must also be ranked	--Recognize significant differences between federal and nonfederal evaluations, re: - audiences/stakeholders - contextual influences - number of data sources - distances - costs - constraints --Standards must be field tested and reviewed at federal level and revised accordingly (need case histories of successes and failures)

Exhibit 2 (continued)

	Pros	Cons	Suggestions
Michael J. Wargo (continued)		--But says informations needs within audiences must also be ranked --Says cost/effectiveness standard overlooks distribution of costs and benefits across different levels of program operation and audiences	--Suggests that A1 give more specific advice about ranking of audiences --Suggests the A4 be reworded and given emphasis on comparative or discrepant nature of most evaluations --Says A8 needs to go further, re: ensuring impact --C1 should recognize multiple clients and calls for a new guideline on this
Joan Boyhoff Baron	--Practical orientation --Potential to improve quality and use --Requirements of careful descriptions good --Cases provide good examples of why program description is important --Good advice, re: searching through existing information --Greatest strength is in sensitizing people to consider all key audiences --Use of a large number of advisors in developing the standards --Open methodological base --Short, readable	--They lack the context needed by the unitiated --Standards mean "how much is enough". These standards aren't standards.	--State officials should have several copies for use and circulation --But should be extended to cover proposals for state funding (object description is also important there). (Describe from perspective of the client) --Give more visibility to the "forming of advisory groups." --Give attention to issues in choosing an internal or external evaluator --Emphasize the importance of determining clients' needs --Add concern for stating the limitations of an evaluation in the final section of every evaluation report --Relable the Standards as principles, practices, or methodological guidelines.

Exhibit 2 (continued)

	Pros	Cons	Suggestions
Steve J. Osterlind	--Delinating criteria encourages integrity of program evaluators (the other side of State's notion that omiting bad practices supports disservice) --Provides direction for inservice programs --Provides key check-points for determining competency of prospective employees --Could foster a community spirit and benchmarks for judging quality evaluations --Responsive to field's need for structure and guidelines as opposed to original methodologies --Functional table is highly useful.		
Kathryn A. Hecht	--Standards intended for a wide audience.	--Style and volume militate against use by lay audiences --But they're too condenced and distilled to be of use by the uninitiated.	--Develop and diversify for use with multiple audiences --Adapt to rural needs and audiences --Develop guidelines on how to integrate into evaluation process --Disseminate through non-print and print media.

Exhibit 2 (continued)

	Pros	Cons	Suggestions
Marcia Linn	--Plethora of sound suggestions.	--Identify important elements more effectively then they characterize trade-offs --Tell more about what not to do than clever things that could be done (e.g., resist pressures and don't missuse power)--May thus inhibit evaluator and reduce his power.	--Augment standards with specific knowledge of the area being investigated and with other conceptual frameworks.
Barbara Searle		--American setting is less centralized, more affluent and utility and propriety standards won't apply as well --No guidelines for rank ordering the Standards --Provincial.	--Interpret them in educational settings, very different from those for which they were developed in order to increase understanding of them and improve usefulness.
Ray Lyons and Bill Rubin	--Cases are extremely informative.		--Development of supplementary book on implementation ideas is a good idea --Sees variety of models and standards to be desireable.

Exhibit 3

An Interpretive Analysis of the Reactions to the Standards

	Pros	Cons	Recommendations
Content of the Standards	--Advice is realistic, practical, usually of admirable temperament, and commanding of adherence --Scope is most thorough and the Standards: - raise a large number of key issues - offer many good suggestions - endorce methodological variety - focus on a variety of evaluation situations - provide a range of interesting and realistic illustrations - confront a range of political realities --Standards is aimed at a varied audience and they: - provide a comprehensive yet concise statement of "good practice" - are short and readable - provide an organization and format which is easy to use - offer a useful functional table of contents	--Some standards are incompatible, but: - trade off problems are not well described - guidance for rank ordering standards is too general --Scope is too limited in: - identification of bad practices in the evaluation establishment - treating needs assessment - suggesting clever tactics - not requiring experimental designs --Standards are difficult to use because they: - are numerous - lack enough context to help the uninitiated - are not ranked for their relevance to licensing and accrediting - are provincial --Standards have limited use at Federal level because they: - assume, unrealistically, that evaluators can quit - fail to recognize and deal with the complexities of team work involved in Federal evaluations - include too few case studies at the Federal level.	--Develop companion volumes to include: - a summplemental book on implementing the Standards - a consumer's guide to program evaluation --Consider adding materials that: - prohibit unrealistic promises - prohibit being indifferent to the concerns of evaluees - promote efforts to understand education - require specific knowledge of area being investigated - encourage use of a variety of conceptual frameworks - call for only valid and defensible designs - project propriety and accuracy standards as the most important in accreditation and licensing - give more visibility to forming of advisory groups - provide guidance for choosing internal and external evaluators - emphasize the importance of determining clients' needs - suggest stating limitations of the evaluation in the final section of the evaluation report - show how to integrate the standards in the evaluation process - show how to employ the standards in licensing

Exhibit 3 (continued)

	Pros	Cons	Recommendations
Content of the Standards (continued)	--Particularly helpful suggestions include: - advice on rank ordering audiences - requirements for careful descriptions - advice on searching through existing information --Standards both emphasize quality evaluation and respond to the field's need for structured and concrete suggestions as opposed to offering vague suggestions for using "original" methods.	--The Audience Identification standard was claimed to be both: - overspecified - underspecified --The Cost Effectiveness standard overlooks the distribution of costs and benefits across different levels of program evaluation.	--Diversify the Standards for use with: - federal audiences - rural audiences --Add specificity in A1 reranking of audiences --Emphasize in A4 that most standards are comparative or discrepant --Expand A8 to promote more impact --C1 should recognize multiple audiences --Consider relabeling the Standards as Principles, Practices, or methodological guidelines
Rationale for Developing and Using Standards	--Case for standards has best been stated by the Committee --The Standards have many potential benefits: - benchmarks for judging evaluations - content for inservice and preservice training - guide for developing standardized tests on evaluation - framework for a competency based approach to accrediting and licensing - framework for developing evaluation contracts	--Committee wasn't sufficiently vigorous in examining the case against the standards --Standards may aid unjust denials of the privilege to practice evaluation --Standards constrain creativity (Stake) --Committee's definition of standards may be to general	--Provide standards by example in meta-evaluations as an alternative to formal standards --Leave room for a variety of models and standards in order to accommodate different settings.

Exhibit 3 (continued)

	Pros	Cons	Recommendations
Rationale for Developing and Using Standards (continued)	- promotion of better conduct and use of evaluation - encouragement for improving the integrety of evaluators --Standards that can't be measured are okay because they arouse people to care about what others admire.		
Structure of the Standard Setting Effort	--The developers of the Standards possessed high credibility --The Committee made good use of a large number of advisors.	--The developers and support groups under-represented certain groups including: - federal agency personnel - contract research organizations and - foreign groups - rural school personnel	--Future developmental efforts should increase the involvement of the previously under-represented groups.
Uses of the Standards			--State agency personnel should maintain a supply of the Standards for distribution to the groups they serve. --Resist uses of the Standards that constrain individual practice if done only to protect the evaluation profession --The Standards should be field tested at the Federal level and in a range of very different settings --The Standards should be disseminated through the use of non-print media.